

Features

Bias and Inclusivity in Metadata

Brian M. Watson

Metadata—as any archivist, librarian, or information professional knows—is critically important. If you can understand and control metadata, you can control the knowledge discipline, even those you do not know anything about, according to Robert D. Montoya, assistant professor of information and library science at Indiana University Bloomington (IUB).

Metadata can be lethal: “We kill people based on metadata,” said the former head of the Central Intelligence Agency and National Security Agency, retired US Air Force General Michael Hayden.¹

*If metadata and cataloging are the “power to name,”
then it is worth asking: Who is doing that naming?*

And yet, metadata can also be liberating—it’s a tool that can be used to take apart the master’s house with the master’s tools.² If metadata and cataloging are the “power to name,” then it is worth asking: Who is doing that naming?³ That query is what Julie Hardesty, metadata analyst and assistant librarian at IUB, has been asking. The answer to that question, I would say, also answers the question of how to build the tools to dismantle the house.

A Ubiquitous Problem

In front of a nearly-packed room in IUB’s Herman B Wells Library and broadcast online, Hardesty began by saying that anyone presenting on these topics “need[s] to be aware of our biases . . . I know some of my own as a white woman from the Midwest but I am still trying to learn.” Arguing that bias in metadata is a massive international problem, she demonstrated it

by pointing out issues at home, in IUB's own catalog system (IUCAT), which assigns the problematic "ethnographic," "Indian," and "Historical" Library of Congress Subject Headings (LCSH) to archival photographs of Native Americans. She strengthened her argument with a case on gender, pointing out how the classification of poets works in IUCAT and in many other digital archives and special collections: it privileges individuals who are "white AND male AND straight AND European AND Christian AND middle-class AND able-bodied AND Anglo."⁴

She concludes that the more marginalized a subject is, the more obscure and restrictive the subject headings on their work are. This is a problem that cuts both ways, Hardesty argues. Not only does it place a majority in charge of naming a minority, it also reinforces that status by limiting the subject's access to their own material, removing it from the center.

Controlled Vocabularies and Classification Schemata

The center, however, cannot hold—and in this case, that is a very good thing indeed. In her research, Hardesty has discovered a variety of projects focused on these problems, which fall broadly into two categories: (1) controlled vocabularies intended for use in archives or library catalogs in addition to established schemes; and (2) alternative classification systems that focus on how items are categorized for placement on the shelf or browsing. During her talk, Hardesty went into further detail on remedial controlled vocabulary projects, including the Indigenous Xwi7xwa Library (University of British Columbia), Atria (Institute on Gender Equality and Women's History, Netherlands), Mashantucket Pequot Thesaurus of American Indian Terminology (University of Washington/Pequot Research Center), Homosaurus (www.homosaurus.org), and the National Indian Law Library in Boulder, Colorado. Some of these are online, some are linked data, some are controlled vocabulary, and some also include classification schemes, but none of them are, as of yet, all of the above.

Hardesty then pivoted to classification schemata, pointing out that even though revision attempts were pioneered by Howard University Librarian Dorothy Porter in the 1930s and '40s, there are still problematic classifications in many archives and libraries, whether they use Dewey Decimal or Library of Congress. To demonstrate this, she again turned to IUCAT, highlighting three recent publications: Nina O'Leary's *Native Enough*, Stephanie J. Waterman et al.'s *Beyond Access: Indigenizing Programs for Native American Student Success*, and Robin Starr Minthorn's *Reclaiming Indigenous Research in Higher Education*. Each of them is classified under LCSH E97 for "History of North America—Indians of North America." All of these items were published in 2018 but are "placed in the past on the shelf," in Hardesty's words.

She offered several alternative cataloging approaches, such as the Brian Deer Classification system (used in Canada) or the Lavender Library LLACE classification system (developed in San Francisco), which is used by Indiana University's LGBTQ+ Center, independently of IUCAT.

This latter classification system, which is completely student-chosen and run, is an “example of what is possible for collection organization when a community historically [minoritized] *becomes* the center and focus of the organization scheme” (my emphasis).

Five Possible Next Steps

Libraries, archives, and museums have a lot of metadata—but almost no resources or tools to use in reviewing, evaluating, or updating them. Hardesty provides a number of possible solutions:

- 1) Provide controlled vocabularies besides LCSH for digital collections,
- 2) Experiment with information retrieval aids using different controlled vocabularies,
- 3) Implement institutional solutions like those recommended by the Digital Library Federation’s Cultural Assessment Working Group and its Inclusive Metadata Task Force,
- 4) Provide contextual information about controlled vocabulary and classification problems to users, and
- 5) “Center the marginalized” to try providing multiple entryways to resources through catalog portals and metadata.

Each of these ideas offers potential tools of power or pathways forward, and should be seriously explored by all those identified, cataloged, and subjected to headings that do not name them. To view Hardesty’s full presentation on “Bias and Inclusivity in Metadata: Awareness and Approaches,” given March 20, 2019, go to <http://hdl.handle.net/2022/22880>. View my live-tweeted thread at <https://bit.ly/2UbtqiG>.

Notes

¹ Farran, Lee. “Ex-NSA Chief: ‘We Kill People Based on Metadata.’” ABC News. ABC News, March 5, 2014.
<https://abcnews.go.com/blogs/headlines/2014/05ex-nsa-chief-we-kill-people-based-on-metadata>.

² Pomerantz, Jeffrey, *Metadata*, The MIT Press Essential Knowledge Series (Cambridge, Massachusetts; London, England: The MIT Press, 2015), 118. Olson, Hope A., “The Power to Name: Representation in Library Catalogs,” *Signs* 26:3 (2001), 639–68, 661.

³ Ibid.

⁴ Hardesty, slide 16. Hardesty provides a diagram from Hope A. Olson, “Mapping Beyond

Dewey's Boundaries: Constructing Classificatory Space for Marginalized Knowledge Domains," *Library Trends* 47:2 (1998), 238.